IMPACT: International Journal of Research in Humanities, Arts and Literature (IMPACT: IJRHAL)
ISSN (P): 2347–4564; ISSN (E): 2321–8878
Vol. 7, Issue 9, Sep 2019, 27–34

APPLICATION OF LOTKA'S LAW ON THE RESEARCH PUBLICATIONS OF THE BOTANICAL SURVEY OF INDIA

Bagavathi A[1] & M.E. Kalyani[2]

[1]Assistant Librarian, The Tamil Nadu Dr. Ambedkar Law University, Chennai, India

[2]Assistant Librarian, Central University of Tamil Nadu, Thiruvarur, India

Received: 19 Aug 2019 **Accepted:** 21 Sep 2019 **Published:** 30 Sep 2019

ABSTRACT

This study includes the analysis of only the journal publications of Botanical Survey of India (BSI) irrespective of its centre's. The authors' productivity pattern in terms of Lotka's law was determined using the Kolmogorov-Smirnov test. In which, the formula $x^n y = c$ of Lotka's was used with the constant determination aspects by the well-known expert on the constant values of the formula.

KEYWORDS: *Bibliometric Study, Lotka's Law, Authorship Pattern, Author's Productivity, BSI*

INTRODUCTION

The article contributions of the 10 well established centre's of Botanical Survey of India (BSI) was taken into an account for this particular study. BSI is a research centre in the field of Plant Taxonomy and its related areas like Systematic Botany, Photo geography of Plants, etc. It comes under Ministry of Environment and Forests of Government of India.

OBJECTIVES

To determine the application of the Lotka's law in the research productions of Botanical Survey of India using the Kolmogorov-Smirnov test.

METHODOLOGY

The data was collected from the annual reports of the BSI and the reference of citation databases like Scopus, Web of Science. As the works on Indian systematic botany developed by BSI were not comprehensively covered by any of the available bibliographical databases, it necessitates relying on the annual reports, project reports, progress reports, etc developed by the institute for the data collection.

REVIEW OF LITERATURE

Ahmed and Rahman (2008) have presented the results of a bibliometric analysis of nutrition literature of Bangladesh. A list of periodical articles on various aspects of nutrition research of Bangladesh published during 1972–2006 was compiled for analysis. A total of 636 articles by 998 authors were identified. The articles were published in 100 local and foreign journals. The five-yearly distribution of nutrition literature proved that there has been a rapid growth of nutrition literature

from 1987 onwards. Lotka's law was found to be applicable to nutrition literature of Bangladesh. They revealed that Bradford-Zipf distribution were also applicable to the literature.

Sevukan and Sharma (2008) made a detailed analysis on biotechnology in central universities of India from 1997 to 2006. The data used for the study were retrieved from the database sources, namely, PubMed, NCBI (National Centre for Biotechnology Information) and Web of Science database—Science Citation Index Expanded (SCIE). The results indicated that the growth of literature in biotechnology has steadily increased from 15 articles in 1997 to 43 articles in 2006; two-authored publications predominated amongst the pattern of authorship; applicability of Lotka's law was validated from the values n = 2.12, C = 0.669, and D = 0.027 obtained using least square method. However, the application of Bradford's law does not fit to the literature analyzed.

Patra and Chand (2007) examined the growth over time of Indian Acquired Immune Deficiency Syndrome (AIDS) research output based on bibliographic data from PubMed and Web of Science. This study's result depended on 2,178 records covering journal articles, case reports and review articles on Indian AIDS research. The study showed that from 1992 onwards, the growth of literature was exponential, and in last two years over 300 articles for a year has been published. Also reflects that the highest number of publications was from the United States followed by the United Kingdom and France. China was ranked below India. A Kolgomorov-Smirnov test showed that literature has not following the original Lotka's distribution.

APPLICATION OF LOTKA'S LAW

The formula of Lotka (1926) that represents the publication productivity of the author is expressed as $x^n y = c$. In which x denotes the number of papers / articles, y states the number of authors contributing the x number of papers/articles and n and c were the constants. He took the data set from the chemical and physical sciences for studying the relationship exists between the number of articles produced and the number of primary authors involved and evolved with a law. This Lotka's law of inverse square states that the authorship productivity held with a general pattern in any body of literature. Various researchers have employed his formula in various fields of subjects. Pao (1985) explained the Lotka's law specifically and expressed to find the value of the two constants n and c for each specific set of data using the concept of the Lotka's law as "the number of authors making a certain number of articles is a fixed ratio to the number of authors publishing a single article". His formulae for calculating the constants n and c were employed in this study. For this testing, only the researchers of BSI were considered irrespective of their position in the journal article contribution. There were 1307 BSI authors who have contributed 6186 of journal articles. The number of articles produced and the frequency of the authors were stated in the Table1 for the application of Lotka's law in the institutional production.

The continuous and almost high frequencies of article publication by the authors were ended at the first 35 terms of Table 1. For the calculation of n, the first 35 pairs of x and y were considered as the value of N. As suggested by Pao (1985) the infinite value in the calculation of c with the formulae will stagnate at the point of P terms and the residual error with P = 20 in the calculation was negligible, so, P was taken with the value 20. But the computed value n is only −1.33 not nearing the value n = 2 as suggested by Lotka. So, the considerations of various levels of pairs from the Table 1 were taken into account to see the n value. For this, N with the number of pairs, 20 has yield n= −1.42, with N = 25 the value of n = −1.46, and even the entire pairs from the Table No. 33 was considered (N = 84) and provided with n = −1.28. Finally, the 50 pairs (N) that constitute the number of authors against the number of articles' generated (Table 1) were considered

for the application of Lotka's Law using the improved method of Pao (1985) in finding the value of the constant n and c. In the data concern the fifty pairs of number of publications produced by the corresponding author were stated in t table 2 and considered for the computation of n using the equation 1.

Using the equation 2, the pendant value of the constant c would easily be calculated as follows. Here the non negative n value is to be considered and P = 20 instead of infinity as stated by Pao (1985) in her theorem of examining the Lotka's Law of author productivity.

$$n = \frac{N\sum XY - \sum X \sum Y}{N\sum X^2 - (\sum X)^2} \quad (1)$$

$$c = \frac{1}{\sum_{x=1}^{P-1} \frac{1}{x^n} + \frac{1}{(n-1)P^{n-1}} + \frac{1}{2P^n} + \frac{n}{24(P-1)^{n+1}}} \quad (2)$$

n=(-559.5952)/379.9354

n = -1.4729

$$c = \frac{1}{\sum_{x=1}^{19} \frac{1}{x^{1.4729}} + \frac{1}{(1.4729-1)P^{1.4729-1}} + \frac{1}{2P^{1.4729}} + \frac{n}{24(20-1)^{1.4729+1}}}$$

$$c = \frac{1}{2.20618 + \frac{1}{(0.4729 * 20^{0.4729})} + \frac{1}{2 * 20^{1.4729}} + \frac{1.4729}{24 * 19^{2.4729}}}$$

$$c = \frac{1}{2.20618 + 0.51283 + 0.006063 + 4.22408}$$

$$c = \frac{1}{2.725115}$$

c= 0.36696

From the above determination, the percentage of authors with only one published work in the theoretical terms was 36.7% while the observed value was 43.4% as shown in the first entry of the table 3.

Here the data considered for testing has exceeded 35 and 50 pairs of data were considered, so the significant level of 0.1 using the following equation provided by Black (2003, p.567) was employed to calculate the critical value.

Critical value = 1.22/√n+1 (3)

Where n is the total of authors considered that comes to 1268 as shown in the Table 3.

Critical value = 1.22 / √1268+1

= 0.034248

The critical value is smaller than the highest value of D as derived above.

Critical Value<D_{max}

Table 1: Number of Publications by Authors for the Application of Lotka's Law

No. of Published works x	No. of Authors Contributed y	No. of Published works x	No. of Authors Produced y
1	550	46	2
2	181	47	3
3	89	48	2
4	55	49	1
5	58	51	1
6	29	52	1
7	32	54	2
8	21	55	1
9	22	57	2
10	18	58	2
11	20	59	1
12	19	60	1
13	9	62	1
14	6	63	1
15	11	64	2
16	11	68	1
17	7	69	1
18	6	70	2
19	7	71	1
20	13	74	2
21	6	77	1
22	5	78	1
23	4	80	1
24	5	81	1
25	4	82	1
26	6	83	1
27	2	87	1
28	3	89	1
29	5	91	1
30	9	95	1
31	8	96	1
32	6	101	1
33	7	103	1
34	6	106	1
35	3	110	1
36	1	123	1
37	1	131	1
39	2	139	1
41	3	150	1
43	1	153	1
44	2	171	1
45	2	174	1

Hypothesis Discussed & Verified

From the above derivation using Kolmogorov–Smirnov test, the null hypothesis revealed that the BSI's author productivity is not different from the theoretical predication of Lotka in his law (1926) about the author productivity($x^n y = c$) was rejected.

Table 2: Calculation of N for the First 50 Pairs of BSI Author's Productivity

X	Y	X = Log X	Y=Log Y	Xy	Xx
1	550	0	2.7404	0	0
2	181	0.3010	2.2577	0.6796	0.0906
3	89	0.4771	1.9494	0.9301	0.2276
4	55	0.6021	1.7404	1.0479	0.3625
5	58	0.6989	1.7634	1.2324	0.4885
6	29	0.7781	1.4624	1.1379	0.6054
7	32	0.8451	1.5052	1.2720	0.7142
8	21	0.9031	1.3222	1.1941	0.8156
9	22	0.9542	1.3424	1.2809	0.9105
10	18	1	1.2553	1.2553	1
11	20	1.0414	1.3010	1.3549	1.0845
12	19	1.0792	1.2788	1.3801	1.1647
13	9	1.1139	0.9542	1.0629	1.2408
14	6	1.1461	0.7782	0.8919	1.3135
15	11	1.1761	1.0414	1.2248	1.3832
16	11	1.2041	1.0414	1.2539	1.4499
17	7	1.2304	0.8451	1.0398	1.5139
18	6	1.2553	0.7782	0.9769	1.5758
19	7	1.2788	0.8451	1.0807	1.6353
20	13	1.3010	1.1139	1.4492	1.6926
21	6	1.3222	0.7782	1.0289	1.7482
22	5	1.3424	0.6989	0.9382	1.8020
23	4	1.3617	0.6021	0.8199	1.8542
24	5	1.3802	0.6989	0.9646	1.9049
25	4	1.3979	0.6021	0.8417	1.9541
26	6	1.4149	0.7782	1.1011	2.0019
27	2	1.4314	0.3010	0.4309	2.0489
28	3	1.4472	0.4771	0.6905	2.0944
29	5	1.4624	0.6989	1.0221	2.1386
30	9	1.4771	0.9542	1.4094	2.1818
31	8	1.4914	0.9031	1.3469	2.2243
32	6	1.5052	0.7782	1.1713	2.2656
33	7	1.5185	0.8451	1.2833	2.3058
34	6	1.5315	0.7782	1.1918	2.3455
35	3	1.5441	0.4771	0.7367	2.3842
36	1	1.5563	0	0	2.4221
37	1	1.5682	0	0	2.4593
39	2	1.5911	0.3010	0.4789	2.5316
41	3	1.6128	0.4771	0.7695	2.6011
43	1	1.6335	0	0	2.6683
44	2	1.6435	0.3010	0.4947	2.7011
45	2	1.6532	0.3010	0.4976	2.7331
46	2	1.6628	0.3010	0.5005	2.7649
47	3	1.6721	0.4771	0.7978	2.7959
48	2	1.6812	0.3010	0.5060	2.8264
49	1	1.6902	0	0	2.8568
51	1	1.7076	0	0	2.9159
52	1	1.7160	0	0	2.9447
54	2	1.7324	0.3010	0.5215	3.0012
55	1	1.7404	0	0	3.0289
Total ∑		**64.8753**	**40.4476**	**41.2891**	**91.7748**

Table 3: K-S Test of Empirical and Theoretical Value of Author Productivity

Empirical value				Theoretical Value		
X	**Y**	$F_{0(Y_{x})}=\frac{y_x}{\sum y_x}$	**Cumulative Expected Value of Authors $F_e(Y_x)=C(1/X^n)$**	**Expected Value of Authors $F_e(Y_x)=C(1/X^n)$**	**Cumulative D = $\sum f_0(y_x) - \sum f_e(y_x)$**	**D = $\sum f_0(y_x) - \sum f_e(y_x)$**
1	550	0.4338	0.4338	0.36696	0.36696	0.06684
2	181	0.1427	0.5765	0.1322	0.49916	0.07734
3	89	0.0702	0.6467	0.072756	0.571916	0.074784
4	55	0.0434	0.6901	0.047626	0.619542	0.070558
5	58	0.0457	0.7358	0.034285	0.653827	0.081973
6	29	0.0229	0.7587	0.026211	0.680038	0.078662
7	32	0.0252	0.7839	0.020887	0.700925	0.082975
8	21	0.0166	0.8005	0.017158	0.718082	0.082418
9	22	0.0174	0.8179	0.014425	0.732507	0.085393
10	18	0.0142	0.8321	0.012351	0.744859	0.087241
11	20	0.0158	0.8479	0.010734	0.755592	0.092308
12	19	0.0150	0.8629	0.009443	0.765035	0.097865
13	9	0.0071	0.8700	0.008393	0.773428	0.096572
14	6	0.0047	0.8747	0.007525	0.780952	0.093748
15	11	0.0087	0.8834	0.006798	0.78775	0.09565
16	11	0.0087	0.8920	0.006181	0.793931	0.098069
17	7	0.0055	0.8976	0.005653	0.799584	0.098016
18	6	0.0047	0.9023	0.005197	0.804781	0.097519
19	7	0.0055	0.9078	0.004799	0.80958	0.09822
20	13	0.0103	0.9181	0.00445	0.814029	0.104071
21	6	0.0047	0.9228	0.004141	0.818171	0.104629
22	5	0.0039	0.9267	0.003867	0.822037	0.104663
23	4	0.0032	0.9299	0.003622	0.825659	0.104241
24	5	0.0039	0.9338	0.003402	0.829061	0.104739
25	4	0.0032	0.9370	0.003203	0.832264	0.104736
26	6	0.0047	0.9417	0.003023	0.835288	0.106412
27	2	0.0016	0.9433	0.00286	0.838148	0.105152
28	3	0.0024	0.9457	0.002711	0.840859	0.104841
29	5	0.0039	0.9496	0.002574	0.843433	0.106167
30	9	0.0071	0.9567	0.002449	0.845882	0.110818
31	8	0.0063	0.9630	0.002333	0.848215	0.114785
32	6	0.0047	0.9677	0.002227	0.850442	0.117258
33	7	0.0055	0.9733	0.002128	0.85257	0.12073
34	6	0.0047	0.9780	0.002037	0.854607	0.123393
35	3	0.0024	0.9804	0.001951	0.856558	0.123842
36	1	0.0008	0.9811	0.001872	0.85843	0.12267
37	1	0.0008	0.9819	0.001798	0.860228	0.121672
39	2	0.0016	0.9835	0.001664	0.861892	0.121608
41	3	0.0024	0.9859	0.001546	0.863438	0.122462
43	1	0.0008	0.9867	0.001441	0.864879	0.121821
44	2	0.0016	0.9882	0.001393	0.866272	0.121928
45	2	0.0016	0.9898	0.001348	0.86762	0.12218
46	2	0.0016	0.9914	0.001305	0.868925	0.122475
47	3	0.0024	0.9938	0.001264	0.870189	0.123611
48	2	0.0016	0.9953	0.001226	0.871414	0.123886
49	1	0.0008	0.9961	0.001189	0.872603	0.123497
51	1	0.0008	0.9969	0.001121	0.873724	0.123176
52	1	0.0008	0.9977	0.001089	0.874813	0.122887
54	2	0.0016	0.9993	0.00103	0.875844	0.123456
55	1	0.0008	1.0001	0.001003	0.876847	0.123253
	1268					**D_{max}= 0.123886**

FINDINGS AND CONCLUSIONS

Based on the dataset of empirical and theoretical values, the application of Lotka's Law was scrutinized. It has been revealed that the BSI's author productivity has not followed the pattern as stated in Lotka's law. Lotka has derived his dataset from Chemical and Physical Sciences which have shown some pattern of relationship while the plant taxonomical dataset of BSI has not confined to the described pattern.

REFERENCES

1. *Ahmed, S. M. Z., & Rahman, Md. A. (2008). Nutrition literature of Bangladesh: A bibliometric study. Malaysian Journal of Library & Information Science, 13 (1), 35–43.*
2. *Bagavathi, A. & Kalyani, M.E. (2018). Growth and Development of BSI in Research Publication on Authorship Pattern. Journal of Advances in Library and Information Science, 7 (4), 326–335.*
3. *Lotka, A. (1926). Statistics: The frequency distribution of scientific productivity. Journal of the Washington Academy of Sciences, 16 (12), 317–323.*
4. *Nicholls, P. T. (1986). Empirical validation of Lotka's law. Information Processing & Management, 22 (5), 417–419.*
5. *Nicholls, P. T. (1989). Bibliometric modeling processes and the empirical validity of Lotka's law. Journal of the American Society for Information Science, 40 (6), 379–385.*
6. *Pao, M. L. (1985). Lotka's law: a testing procedure. Information Processing & Management, 21 (4), 305–320.*
7. *Pao, M. L. (1986). An empirical examination of Lotka's law. Journal of the American Society for Information Science, 37 (1), 26–33.*
8. *Patra, S. K., & Chand, P. (2007). HIV/AIDS research in India: A bibliometric study. Library & Information Science Research, 29(1), 124–134.*
9. *Sevukan, R., & Sharma, J. (2008). Bibliometric analysis of research output of Biotechnology faculties in some Indian Central Universities. DESIDOC Journal of Library & Information Technology, 28(6), 11–20.*
10. *Sobrino, M. I. M., Caldes, A. I. P., Guerrero, A. P. (2008). Lotka law applied to the scientific production of information science area. Brazilian Journal of Information Science, 2(1), 16–30.*
11. *Sudhier, K. G. P. (2013). Lotka's law and pattern of author productivity in the area of Physics research. DESIDOC Journal of Library & Information Technology, 33(6), 457–464.*